Global Approaches to Chromatin

Bradley E. Bernstein1,2,3 and Stuart L. Schreiber1
1Department of Chemistry and Chemical Biology
Bauer Center for Genomics Research
and the Howard Hughes Medical Institute
Harvard University
Cambridge, Massachusetts 02138
2Department of Pathology
Brigham and Women’s Hospital
Boston, Massachusetts 02115
3Correspondence: bernstei@slsiris.harvard.edu

Eukaryotic genomes, which comprise as many as 3 billion DNA bases, are packaged into a higher-ordered structure called chromatin. Over the last decade, biochemical and genetic analyses have led to an almost revolutionary understanding of chromatin as a fundamental regulator of genome function rather than as a structural scaffold. However, studies that focus on specific genes, proteins, and histone modifications are limited in their ability to describe comprehensively the structure and function of chromatin. Systematic approaches, reviewed here, are beginning to achieve this kind of global perspective.

Systematic approaches are profoundly influencing our understanding of biology and medicine. These approaches, which rely on sophisticated technologies in chemistry and biology, have the potential to comprehensively interrogate cellular states and perturbations. Rather than focusing on specific genes, proteins, or pathways, these studies aim to provide a global perspective that reflects many or all aspects of cellular function. Transcriptional profiling, in which mRNA expression levels are simultaneously measured for many or all genes in an organism, is the most accessible of these techniques. For example, studies that classify tumors on the basis of mRNA expression profiles have led to the identification of new cancer subtypes and are beginning to provide important clinical information such as prognosis and drug sensitivity [1].

Proteomic analysis, though less accessible than nucleic acid technologies, is also becoming feasible as a result of sophisticated mass-spectrometry instrumentation, bioinformatic algorithms, and the availability of whole genome sequences. Applications of this technology include the determination of cellular protein profiles (conceptually analogous to transcriptional profiling) [2] as well as the comprehensive annotation of the protein components of cell organelles [3] and protein complexes [4]. In complementary tour de force studies, two groups have combined protein purification and mass spectrometry (LC-MS) to identify and determine comprehensively the content of nearly every protein complex in yeast [5,6]. The wealth of informatics data generated by these studies should prompt new discoveries in many areas of biology.

A third area in which systematic approaches are yielding biological insight involves the use of small molecule modulators of biologic function. Diversity-oriented synthesis (DOS) and high-throughput screening (HTS) are now being applied toward the study of fundamental biology. In an example that illustrates the potential of these techniques, a small molecule modulator of the Ure2 signaling protein was discovered by DOS and the use of small molecule microarrays and used to dissect the nutrient response signaling network [7].

We review systematic approaches in biology specifically as they have been applied toward the study of chromatin, a complex field that is highly amenable to these technologies. Experiments and gained insights are presented, with the thesis that these approaches are yielding a more comprehensive understanding of chromatin structure and function unachievable by conventional methodologies.

Chromatin Overview

The extraordinary size and complexity of eukaryotic genomes pose a significant challenge to the cellular machinery, which must maintain, replicate, and transcribe their. This size and complexity also poses a significant challenge to scientists who seek to understand it. The cell’s solution is to package its genome into a higher-ordered, highly regulated structure called chromatin (Figure 1). The fundamental unit of chromatin is the nucleosome: 146 bps of DNA wrapped around an octamer of histone proteins. Histones are subject to a vast array of posttranslational modifications primarily within their amino-terminal tails. Enzymes that acetylate (HATs), deacetylate (HDACs), phosphorylate (HKs), and methylate (HMTs) histones are recruited to specific sites within the genome where they influence transcription, DNA repair, replication, and other processes (for review see [8]). An increasing body of literature indicates that these modifications act by recruiting additional regulatory enzymes that effect chromatin structure [9-11].

Researchers have used a variety of techniques to document fundamental roles of histone modifications in DNA repair, replication, and transcription. Chemical analyses have identified the locations and natures of modifications (e.g., acetylation of specific lysines in the N-terminal tails of histones H3 and H4), biochemical analyses have identified histone-modifying enzymes (e.g., Sir2 is an HDAC [12]), and genetic analyses have documented function for modifying enzymes and other chromatin proteins (e.g., Sir2 is required for proper repression or “silencing” at telomeres [13]). A particularly powerful approach that combines elements of molecular biology and biochemistry is chromatin immunoprecipitation (chip). In chip, fragmented chromatin prepared from cross-linked cells is immunoprecipitated using antibodies specific for posttranslationally modified histones (e.g., H3 acetylated at Lys9 and Lys14). Since the histones are cross-linked to DNA, this procedure enriches for DNA associated in vivo with histones that
Figure 1. Elements of Chromatin Regulation

Eukaryotic DNA and histone proteins are packaged into nucleosomes. Posttranslational modifications to the histone tails, catalyzed by several enzyme families, are central to chromatin regulation. These modifications act by recruiting downstream regulatory proteins that contain acetyl-lysine binding bromodomains or methyl-lysine binding chromodomains. Systematic approaches are playing important roles in the pursuit of a more complete understanding of chromatin structure and function.

Phenotypic Analyses of Yeast Chromatin Mutants

Histone proteins, and in particular their highly modified amino-terminal tails, are highly conserved throughout the eukaryotic kingdom. Hence, Saccharomyces cerevisiae (budding yeast) is an ideal model system that has been widely exploited for the study of chromatin. The sequencing of the S. cerevisiae genome, completed in 1996 [15], has enabled the development of array technologies capable of simultaneously measuring mRNA levels for essentially all genes in the organism. Most of the transcriptional profiling studies that have been carried out since have been based on one of two platforms. In the first, differentially labeled experimental (mutant) and control (wild-type) cDNAs (generated by reverse transcribing mRNA) are combined and hybridized to a single cDNA microarray that contains nearly all yeast open reading frames [16]. In the second, a single sample of labeled RNA is hybridized to a high-density array of oligonucleotides that complement specific sequences within each yeast gene [17]. Datasets collected with cDNA microarrays contain values for each yeast gene that reflect the ratio between experimental and control mRNA levels. Datasets collected by high-density arrays contain values for each yeast gene that reflect absolute mRNA levels in a single experimental or control strain.

In one of the first global chromatin studies, Wyrik and colleagues used high-density arrays to analyze transcription genome-wide in yeast lacking components of the repressive SIR complex [18]. Conventional studies had documented a role for the Sir proteins in the repression or “silencing” of reporter genes introduced near the ends of specific yeast chromosomes (“telomeres”) [13]. Global analysis extended this observation, documenting a role for these proteins in the repression of endogenous genes at the ends of all 16 chromosomes in yeast. The comprehensive nature of this study also enabled the authors to characterize the topology of telomeric silencing. They found that wild-type silencing extends at least 20 kbp from chromosome ends; that is, genes up to this distance are, on average, expressed at 5-fold lower levels than nontelomeric genes. Interestingly, the Sir proteins only silenced genes within 6–8 kbp of chromosome ends, suggesting a role for other factors in repression beyond this range. Later, global analyses would reveal that the histone methyltransferase Set1 silences genes located up to ∼20 kbp from chromosome ends [19]. This suggested that Set1 acts in a pathway distinct from the Sir proteins to maintain silencing, as is consistent with genetic analyses [20].

Global transcriptional analyses of the transcription activating complexes TFIIID and SAGA have also proved illuminating. In contrast to the SIR complex, which contains an HDAC (Sir2), these complexes contain HATs (TAF1145 and Gcn5 in TFIIID and SAGA, respectively). Here again, high-density arrays were used to analyze transcription in yeast lacking subunits either specific to one or common to both of these complexes [21]. Both complexes were found to have broad regulatory functions: ∼30% or ∼12% of the genome is dependent on subunits specific to TFIIID or SAGA, respectively. They were also found to act redundantly in that a full ∼70%
of the genome is dependent on subunits common to both complexes. An in-depth analysis of the histone acetyltransferase components revealed that for a large fraction of the genome, transcriptional changes are only evident in yeast lacking both the TAF145 and the Gcn5 acetyltransferases. Hence, although these histone modifying enzymes act broadly throughout the genome, each enzyme is able to compensate for the other to maintain transcriptional activity.

Comprehensive transcriptional analyses of the HDAC repressors in yeast have been reported [22]. This study used a combination of high-density arrays and cDNA microarrays to examine mutants for the different HDAC genes. They revealed that the predominant zinc-dependent HDACs, Rpd3 and Hda1, regulate distinct gene targets and functional pathways (cell cycle and carbohydrate metabolism, respectively). A limitation of this and other mutant profiling analyses is that steady-state yeast mutants exhibit many downstream effects that make it difficult to differentiate primary (direct) targets from secondary ones. Time-dependent profiling studies using selective small molecule inhibitors have the potential to overcome this limitation because direct targets should be regulated immediately upon drug treatment. The potent HDAC inhibitor trichostatin was used in the systematic HDAC study for this purpose. Although HDACs are repressors, Rpd3 has previously been found to activate reporter genes inserted at telomeres. Indeed, the transcriptional profile of yeast lacking Rpd3 reveals that this HDAC activates, directly or indirectly, 40% of genes within 20 kbp of the chromosome ends. Like RPD3 deletion, trichostatin treatment regulates many telomeric genes. However, most of these genes are regulated slowly by this inhibitor, suggesting that this paradoxical effect of Rpd3 is, at least in part, mediated indirectly. The utility of trichostatin is limited as this small molecule inhibits Rpd3, Hda1, and other zinc-dependent HDACs. Thus, transcriptional changes could not be specifically attributed to a single enzyme. Isoenzyme-specific inhibitors will be extremely useful in the systematic dissection of HDAC function.

More broadly, we expect that small molecule modulators of protein function will prove particularly valuable to the study of chromatin. We envision two general approaches that exploit the rapid, time-dependent effects of these reagents. First, small molecule inhibitors of histone-modifying enzymes will be used to identify direct targets of these enzymes and to characterize their regulatory effects. Second, small molecules that induce robust changes in the activity of specific genes or gene classes will be used to investigate the mechanisms by which these genes or classes are regulated. For example, various chromatin mutants could be examined for the ability or inability to regulate specific targets in response to a chemical perturbation. This “chemical epistasis” analysis has been used to great effect in the partitioning of the transcriptional wiring of the nutrient response network [23] and should be an especially powerful tool for dissecting chromatin function.

Genome-Wide chIP Studies
Currently, the best approach for systematic identification of direct targets of chromatin-modifying enzymes is a technique that combines chIP and microarray technology. This methodology actually localizes DNA binding proteins (or modified histones) to specific regions of the genome. The protocol begins with cross-linking, chromatin fragmentation, immunoprecipitation, and DNA purification as in conventional chIP. Then, instead of quantitative PCR, either random-primer [24,25] or ligation-mediated PCR [26] is used to amplify and differentially label the enriched DNA and an unenriched control sample. Finally, both samples are combined and hybridized against microarrays containing both the open reading frames (ORFs) and the intergenic regions (INTs) which contain the gene promoters. Initial studies targeting the Gal4, Ste12, SBF, and MBF transcription factors validated genome-wide chIP by confirming (and extending) prior findings [24,26]. Since then, several studies have used this approach to improve our understanding of chromatin.

For example, genome-wide acetylation and methylation patterns have been determined in yeast using this technique (Figure 2) [19]. These data proved particularly illuminating when examined in the context of transcriptional profiling datasets. Specifically, global measures of acetylation and H3 Lys4 methylation were compared to a global measure of transcriptional activity determined by high-density arrays. Consistent with prior studies documenting a role for acetylation in gene promoters, transcriptional activity correlated globally with acetylation of histones in gene promoters. Like acetylation, H3 Lys4 methylation had previously been linked to transcriptional activity. However, no correlation was observed between transcriptional activity and Lys4 methylation in promoters. Instead, transcriptional activity correlated with Lys4 methylation in coding regions. To probe the functional significance of coding region methylation, a yeast mutant lacking the sole Lys4 methyltransferase, Set1 [27,28], was examined. This mutant exhibited a general defect in the transcription of active, coding region methylated genes, indicating that a key function of Lys4 methylation is to facilitate transcription. This effect is likely mediated via influences on histone acetylation and/or transcriptional elongation [19].

Although Lys4 methylation has generally been associated with transcription, Set1 is required to maintain repression at the silent loci in yeast (HML, RDA and telomeres). Furthermore, conventional chIP studies revealed methylation of histones in the rDNA, leading to speculation that Lys4 methylation might, in certain situations, be repressive [20,28]. Global analysis, however, showed that histones at silent loci actually exhibit very little methylation relative to the genome average [19]. Hence, Set1 appears to exert its repressive influence on these loci indirectly. In sum, the identification of coding region Lys4 methylation as a facilitator of transcription and the resolution of an apparent Lys4-silencing paradox represent important contributions of global analysis not forthcoming by conventional approaches.

In a concurrent study, Grunstein and colleagues used complementary approaches to characterize the functions of the yeast HDACs. Noting the limitations of steady-state transcriptional profiles of HDAC mutants, these investigators reasoned that a comprehensive analysis should include two additional elements: ge-
Figure 2. Procedure for Genome-Wide Analysis of Histone H3 Acetylation
For detailed methods see [19,24-26,52,53].

Figure 3. Histone-Modifying Enzymes Act on Distinct Regions of the Genome
Genome-wide chIP studies of histone modifications and modifying enzymes reveal a division of labor among HDACs: (A) Sir2 acts at telomeres [25], Hda1 acts on subtelomeric regions [29], Rpd3 is excluded from telomeric and subtelomeric regions [30], and (B) Hos1/Hos3 deacetylate the rDNA locus [29]. The Lys4 HMT Set1 methylates histones in transcribed regions and is relatively inactive at telomeres, subtelomeres, and the rDNA [19]. These findings suggest that paradoxical effects of Rpd3 and Set1 at telomeres and rDNA are mediated indirectly (see text).

Screening Studies in Chromatin
Systematic approaches can also be used as screens to identify genes, proteins, or small molecules with desired characteristics. For example, a transcriptional profile of a yeast mutant lacking the Set1 protein was used as a preliminary screen to identify targets of this HMT. Potential targets were collated from a list of genes regulated in the mutant and examined by conventional chIP. The PPH3 gene requires Set1 for expression and is associated with histones methylated at Lys4, indicating a role for this modification in its activation [33]. Further analysis at this locus using specialized antibody reagents led to a more precise understanding of the relationship between histone methylation states and transcriptional status.

Another screening study used mass spectrometry to probe the function of Lys4 methylation [34,35]. The HP1 chromodomain had previously been found to bind histone H3 tail methylated at Lys9 [9,10]. The existence of other chromatin factors that specifically interact with H3 tail when methylated (or unmethylated) at Lys4 was hypothesized by analogy. Such factors were sought in differential pull-down experiments. Specifically, pro-
In other areas of biology, are sure to be useful tools.

regions deacetylated by the yeast HDACs Rpd3 and protein array methods [3,50,51], already being applied
regions of the genome that are Lys4-methylated and vances will be made, though mass spectrometry and
revealed a nearly complete lack of overlap between ciated factors. It is premature to predict how these ad-
part, by precluding histone deacetylation. Further sup-
39]. Hence, Lys4 may facilitate transcription, at least in A true global understanding of chromatin awaits also
methylated or Lys9-methylated, but not Lys4-methyl-
protein species that specifically associated with un-
motivated. A truly comprehensive understanding of chro-
peptides corresponding to unmethylated H3 tail or to studies have begun to illuminate the genome-wide func-
tions of chromatin remodeling enzymes in yeast [44-47].
Importantly, technical and biological lessons learned in
this organism will facilitate analogous studies in higher organisms. Although mammalian genomes are orders
of magnitude more complex, preliminary studies have
demonstrated the feasibility of these approaches in hu-
man tissue culture cells [48,49].

Finally, screens for small molecule modulators of chromatin enzymes and associated pathways have
been performed. These kinds of molecules are of inter-
est not only for their potential clinical value [40], but also for their ability to further our understanding of chromatin
structure and function. To reach their full potential in instructing biology, inhibitors should be specific to an
enzyme class (e.g., class I HDACs) or, ideally, to a single iso-
zyme (e.g., HDAC1). HDAC-biased libraries have been
developed [41] and screened with precisely this aim in mind (S. Haggarty, K. Koeller, J. Wong, and S.L.S., unpublished data). Such molecules will facilitate the functional characterization of the different HDAC iso-
zymes in yeast, and, importantly, in other eukaryotes that are less genetically tractable.

In addition to identifying modulators of known pro-
teins (for use in reverse chemical genetics), chromatin researchers are screening small molecules for the ability
to confer particular phenotypes in cells (forward chemi-
cal genetics). For example, two groups screened small
molecules for their ability to disrupt telomeric silencing
in yeast. Grozinger and colleagues identified a class of a hydroxyl-napthaldehyde containing silencing inhibitors
and subsequently found these to be specific inhibitors of yeast Sir2 and its human homolog SIRT2 [42].
Independently, Bedalov and colleagues identified a structur-
ally related, but more potent analog they termed “splito-
imic” (Figure 4) [43]. Profiling studies of yeast treated
with splitomicin and the protein synthesis inhibitor cycloheximide demonstrated that Sir2 is unable to acti-
ivate transcription directly. Furthermore, by treating G1-
Figure 4. Inhibitors of the NAD-Dependent HDAC Sir2
Structurally related small molecules were independently discovered in cell-based screens of different chemical libraries [42,43].

Acknowledgments
We would like to thank Emily Humphrey, Aly Shamji, Rachel Erlich, Marty Burke, and other members of the Schreiber laboratory and the Bauer Center for Genomics Research for helpful discussions and feedback on this manuscript. B.E.B. is supported by a K08 Development Award from the National Cancer Institute. S.L.S. is an Investigator at the Howard Hughes Medical Institute.

References


